DOCUMENT RESUME

ED 462 905 HE 034 782

| | |
|---|---|
| AUTHOR | Hoyt, Donald P.; Pallett, William H. |
| TITLE | Appraising Teaching Effectiveness: Beyond Student Ratings. IDEA Paper. |
| INSTITUTION | Kansas State Univ., Manhattan. IDEA Center. |
| REPORT NO | IDEA-36 |
| PUB DATE | 1999-11-00 |
| NOTE | 8p. |
| AVAILABLE FROM | Kansas State University, IDEA Center, Inc., 1615 Anderson Avenue, Manhattan, KS 66502-4073. Tel: 800-255-2757 (Toll Free); Tel: 785-532-5970; Fax: 785-532-5637; e-mail: IDEA@ksu.edu. |
| PUB TYPE | Guides - Non-Classroom (055) |
| EDRS PRICE | |
| DESCRIPTORS | *College Faculty; Evaluation Methods; Higher Education; *Instructional Effectiveness; *Student Evaluation of Teacher Performance; *Teacher Effectiveness; *Teacher Evaluation |

ABSTRACT

Evaluating faculty effectiveness is important in institutions of higher education. Although evaluation is inherently threatening to most faculty members, the vast majority take their assignments seriously and want to conduct them as effectively as possible. Assessing faculty performance is a complex and time-consuming process. If it is done poorly or insensitively, it can have an adverse effect on institutional quality. Whether or not individual institutions elect to commit the resources required for valid evaluations depends on the degree to which they agree with these propositions: (1) all members of the institution should be accountable for their activities and performance; (2) the conduct and use of credible evaluation programs have an important influence on the welfare and future excellence of the individual, the department, and the institution; and (3) when improvement efforts are supported by institutional policy and guided by comprehensive and valid appraisals of current functioning, the well-being of the individual and the institution are affected positively. (Contains 18 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 462 905

# Appraising Teaching Effectiveness: Beyond Student Ratings

By Donald P. Hoyt, William H. Pallett

IDEA Paper #36
**IDEA CENTER**

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

W. Cashin

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE



HE033 182

IDEA PAPER #36

NOVEMBER 1999

IDEA CENTER

# Appraising Teaching Effectiveness: Beyond Student Ratings

Donald P. Hoyt and William H. Pallett
IDEA Center

Evaluating faculty effectiveness is important in nearly every institution of higher education. Assessing the effectiveness with which various functions are performed is essential to a variety of important administration recommendations and decisions. It also provides feedback which influences the faculty member's self-image and professional satisfaction. And it establishes a climate which communicates the institution's commitment to professional improvement and confidence that every faculty member will make a valuable contribution to the achievement of shared goals.

There are two types of contributions faculty members make to the programs of a department/institution—*indirect* and *direct*. Indirect contributions, while not impacting directly upon the achievement of a program's objectives (principally *student learning*, in the case of instruction; *new insights and knowledge*, in the case of research; *assistance to clients* in the case of service) make a difference to the program's success by affecting the environment of the department/division, the appropriateness and quality of its plans, and the attitudes and skills of other members. Direct contributions are those in which the achievement of program goals is impacted by the individual's personal intervention or involvement.

In many institutions, research and service programs are vitally important. Assessing a faculty member's contributions to each constitutes a serious challenge. However, this paper is concerned only with the instructional program, a focus which is central to the mission of almost every institution of higher education.

## Assessing Instructional Effectiveness

Direct contributions to the instructional program. Most institutions employ a "student rating" system to assist in the evaluation of instruction. Obtaining student feedback is not only a relatively simple procedure but also is one which has considerable credibility for several reasons. (1) Input is received from a number of raters so that *reliability* is usually quite high. (2) Ratings are made by those who have consistently observed the teacher over many hours, so that they are based on *representative behavior*. (3) Observations about student learning, the object of instruction, are made by those who have been personally affected and therefore have high *face validity*. An enormous volume of research supports the credibility and validity of student ratings (Aleamoni, 1981; Cashin, 1995; Braskamp & Ory, 1994).

On the other hand, student rating systems have several important limitations. (1) Some of them are *poorly constructed* (ask questions about matters which are unrelated to student learning; employ words with unclear meanings; double-barreled questions; response alternatives which fail to exhaust the possibilities; etc.). (2) In some instances, administrative procedures have not been standardized, so that results are *not comparable* from one faculty member to the next. (3) Some systems fail to take into account *extraneous influences* (factors which influence ratings but which are beyond the instructor's control, such as class size, student academic motivation, or course/disciplinary difficulty). (4) *Technical and statistical support* is lacking for some systems, so that interpretation of results is problematic.

Even when these potential difficulties are adequately addressed, authorities are agreed that there are a number of important matters related to teaching effectiveness for which students are unqualified to provide valid reports. Cashin (1989) lists 26 specific considerations which he regards as relevant to instructional effectiveness; students are unqualified to provide valid observations for 11 of these, including an array of factors related to subject matter mastery, course design, and curriculum development. Similarly, Cohen and McKeachie (1980) identified 10 criteria of teaching effectiveness which colleagues, but not students, could assess, two of which describe *indirect* influences (commitment to teaching and support for departmental efforts). Keig and Waggoner (1994) synthesized the Cohen and McKeachie criteria into three general features of teaching effectiveness which students are unable to judge validly: (1) the goals, content, and organization of course design, (2) methods and materials used in delivery, and (3) evaluation of student work, including grading practices.

There is a general consensus that students are unable to judge such vital matters as currency of course content or the degree to which it provides a representative (as opposed to biased) view of the subject matter. Nor can they judge the clarity, comprehensiveness, or realism of objectives, the degree to which readings and other assignments are balanced and appropriate, the validity of procedures for assessing student achievement, or the degree to which grading standards are in line with the department's or institution's expectations or policies.

HE031782

How should the gaps created by shortcomings in student ratings be closed? A wide variety of suggestions have been made. Most frequently cited are self-reports, colleague ratings, and ratings by department heads/chairs.

Seldin (1999) has recently reviewed the value and limitations of self-reports. Clearly, self-interest limits the use of these in the evaluation of teaching effectiveness for administrative purposes. But a reflective analysis on the part of the instructor can be instrumental in promoting instructional improvement (Braskamp & Ory, 1994). In addition, the instructor is the only person who can supply certain kinds of information needed by those charged with making such evaluations, including information about course objectives; readings, assignments, and other learning activities; the creation of instructional materials or learning opportunities; procedures for appraising student achievement; results of, and course modifications based on, classroom research and other faculty efforts directed to improving instructional skills. Such reports are commonly made through the faculty member's annual report of instructional efforts; an illustrative outline for making such reports is available on the Center's web page *(www.idea.ksu.edu)* or, in hard copy form, from the Center; request Appendix A. The annual report can be useful in developing the faculty member's *teaching portfolio* (Seldin, 1993; 1997; Zubizarreta, 1999), a device for organizing relevant information for both *appraisal* and *improvement* purposes.

Most of the crucial features of instruction which students are not qualified to judge can, under certain circumstances, be assessed by faculty colleagues. How this is best done is controversial. Centra (1993) has summarized research related to peer classroom observation and concluded that, as currently practiced, such observations are neither reliable nor valid. On the other hand, DeZure (1999) has identified seven steps which can be taken to overcome these shortcomings, including the use of multiple observations and observers, the training of observers, and the employment of a validated observation instrument. Examples of such instruments are found in Seldin's (1999) recent book; observers are expected to rate such factors as knowledge of the subject, enthusiasm, sensitivity to the class' level of knowledge, preparation and organization, and clarity of presentation[1].

Although colleagues may be able to assess such factors with acceptable reliability and validity, impositions on faculty time makes such a process unrealistic on many campuses. DeZure points out that, besides a 2-4 hour training commitment, colleagues must be prepared to spend about four hours in each observation (pre-observation meeting of 30 minutes, 30 minutes to review materials, 75 minutes to observe, 60 minutes to prepare a joint report, and 45 minutes in a post-meeting with the instructor).

There is reason to believe that such an extensive time commitment may be necessary when classroom observation is geared to *instructional improvement* (see **Summative and Formative Purposes, pp. 3–4**). However, when the purpose is primarily to arrive at a summary estimate of teaching effectiveness, the rating of relevant materials is an attractive alternative to classroom visitation. In general, three raters are asked to make independent judgments based on these materials and then, through discussion, arrive at a consensus[2]. If these ratings are guided by a carefully developed instrument, the consensus rating will usually possess satisfactory reliability and validity. A process for collecting such ratings, and a form for expediting this process, are available on the Center's web site; for hard copy, request Appendix B from the Center.

The department or division head/chair is responsible for gathering and synthesizing all evaluation information. Time constraints imposed by classroom visitations or an in-depth review of instructional materials will generally prevent this administrator from making a personal assessment of *direct* contributions to the instructional program. But he/she can be an important source of information in assessing both the faculty member's *scholarly excellence* and his/her *indirect* contributions (see the following section). The head/chair is also in a better position than anyone else to judge the degree of professional responsibility exhibited by the faculty member through such activities as submitting grades, communicating text/library needs, pursuing professional development opportunities, conducting classroom research, and developing innovative instructional materials or opportunities, all of which are relevant to the achievement of excellence in the instructional program.

A form for guiding the head/chair's review of instruction is included on the Center's web site; for a hard copy, request "Appendix D" from the Center.

Indirect contributions to the instructional program. There is general agreement that a department's/institution's "productivity" (success in achieving its goals and objectives) is affected by such matters as "faculty morale", "collegiality", and "faculty vitality". However, little attention has been paid to the responsibility of individual faculty members for contributing to these "facilitative features" of the academic environment.

In terms of the instructional program, there are at least three types of indirect contributions which individual faculty members can make.

1. The *general learning environment.* Through their social and professional demeanor, faculty members influence the "climate" of the department—its openness, objectivity, tolerance of ambiguity, etc. In their interaction with departmental colleagues, faculty members who share teaching ideas, express interest in the instructional work/concerns of others, or who regularly model intellectual curiosity and excitement make contributions to the learning environment which almost certainly will indirectly affect student learning in a positive manner.

2. *Course and curricular development.* Course revision/updating and the development of new instructional materials or learning aids are two ways of making indirect contributions to student learning. Keeping abreast of instructional/curricular innovations and sharing these with colleagues can make a similar contribution. Likewise, faculty members who are actively involved in the curriculum revision process and who explore with colleagues ways to improve the integration/articulation among specific courses can be expected to have a positive impact on student achievement.

---

[1]Some of these (e. g., enthusiasm; clarity of presentation) represent factors which students are able to assess with reasonable accuracy and may therefore be excluded from colleague ratings (unless there is a need or desire to confirm student ratings).

[2]Small campuses often employ only one or two faculty members in each discipline, cooperative arrangements with other institutions may make it feasible to obtain ratings from those in the same discipline. In such instances, consensus can be achieved through mail or telephone consultation.

3. *Improving teaching effectiveness of others.* Indirect contributions to student learning are made when faculty members consult with each other on teaching methods or strategies or exchange classroom visits for purposes of offering constructive critiques. Similar contributions are made by sharing with colleagues information about an innovative assessment method or a new experiential component to a course. In departments which employ graduate teaching assistants or temporary faculty, indirect contributions can be made by those who offer advice or other assistance to their less experienced colleagues.

In most cases, the academic department chair/head is in a good position to judge the indirect contributions of individual faculty members. But it is desirable to obtain additional evidence by polling the teaching faculty. While not every participant will be able to judge the contributions of every faculty member, it is important that all who are able to make relevant observations be asked to do so.

A form for collecting such views is included on the Center's web site; for a hard copy, request Appendix C from the Center.

## Summative and Formative Purposes

Authorities in educational evaluation have traditionally distinguished between *summative* and *formative* evaluation. The former is done as an aid to administrative decision-making; the latter focuses on using evaluative information to improve performance.

Administrative recommendations/decisions. There are four inter-related administrative decisions or recommendations for which conclusions about the individual's teaching effectiveness are important.

1. On the assumption that those who are most successful in their assignments should receive the largest salary increments, many institutions have adopted a "merit increase" policy. At such institutions, the faculty member's merit evaluation is based in part upon the evaluation of his/her contribution to the instructional program.[3]

2. For non-tenured faculty, decisions must be made annually with respect to retention. Unless the evaluation of teaching effectiveness suggests that the faculty member meets, or will be able to meet, the standards for acquiring tenure, it is not in the best interest of the institution to retain the faculty member; in such instances, retention is not in the best interests of the faculty member either, although this may be difficult to accept.

3. At most institutions, a decision about awarding tenure must be made after a period of time (usually six years). Such a decision has critical implications for both the department's fiscal status and its long term quality. Because instruction is a vital function, the tenure policy at most institutions is intended to insure that it will be continuously performed at a high level of quality.[4]

4. Most institutions accord a "rank" to faculty members. Presumably, those of higher rank are more valuable to the institution (contribute more to the achievement of its mission) than those of lower rank. Policies with respect to rank often involve considerations beyond an assessment of effectiveness in performing instructional, research, and service assignments[5]. Nonetheless, those at a given rank are expected to conduct their assignments with acceptable levels of success. Therefore, evaluation of professional effectiveness is essential.

Evaluations whose purpose is exclusively "summative" (to aid in making administrative recommendations or decisions) should focus on *outcomes*. A central question is, "How successfully were the objectives of the course addressed?"[6] Other outcomes may also be relevant; i. e., the production of innovative learning materials; the introduction of creative projects or other extra-class assignments; the conduct of classroom research which tests instructional hypotheses; etc. While all evaluations should be conducted carefully and thoroughly, those whose purpose is summative ask only, "To what degree did the instructor have a favorable impact on outcomes pertinent to the goals of the instructional program?"; they need not gather information about techniques, strategies, or plans which were responsible for this effect.

Recommendations based upon summative evaluations are extremely serious. They affect both the lives of individual faculty members and the welfare of the department (and, ultimately, the institution). Therefore, they should be done with great care. Those with respect to rank and tenure are especially vital since it is nearly impossible to correct a poor decision.

Special care should be taken to ensure that the summative evaluations used to support such decisions are based on a representative and comprehensive review of the faculty member's contributions. In terms of the instructional function, this means that (a) evidence of effectiveness should be available for every course the faculty member has taught (although not necessarily for each term), (b) the evaluation should be based on a cumulative record of the faculty member's teaching effectiveness (usually involving a minimum of six classes); and (c) trends in teaching effectiveness (improvement, steady-state, decline) can be detected.

Improving performance. In contrast to the limited focus of summative evaluation, *formative* evaluation requires much more information. Not only is it necessary to assess the instructor's impact (positive or negative) on outcomes, but also to examine characteristics of the instructor which account for this impact.

It is not necessary to obtain formative evaluations of every course each time it is taught. In fact, experience suggests that instructional improvement is best facilitated by concentrating not only on one course at a time but also on a limited number

---

[3]The amount of influence which this assessment has on the overall merit evaluation is usually determined by a statement describing all faculty responsibilities and the relative importance of each. In some departments, the same relative importance of teaching, research, and service is assigned to every faculty member; in others, these assignments differ among faculty.

[4]Tenure criteria and standards vary among institutions. Almost all require an evaluation of how effectively the faculty member contributes to the department's programs. Many also include assessments of matters not considered in this paper, such as contribution to departmental diversity, cohesion, and collegiality, or evidence that the faculty member will continue to grow in vitality and professional sophistication.

[5]Traditionally, the level of difficulty or complexity of professional assignments, whether in teaching, research, or service, differentiates among ranks. In the instructional area, those of highest rank are usually expected to be the most versatile in terms of the variety of courses they can offer; frequently, they provide advanced and specialized courses which those of lower rank are not yet qualified to teach.

[6]Student learning is affected by many matters, including the motivation of enrollees to learn, the adequacy of their background, and their academic habits and skills. Since faculty evaluation is concerned with the contribution the faculty member made to student learning, it is desirable to exclude (take into account) the contribution made by such "extraneous" influences. For this reason, the IDEA system provides "adjusted" ratings.



BEST COPY AVAILABLE

of features of the course (e. g., examinations; gaining student involvement; selection and assignment of course projects; etc.).

A relatively sophisticated aid to instructional improvement is offered by the IDEA "student rating" program. The Center's research program established the relationship between 20 specific instructional approaches, chosen in part to represent seven principles of good practice (Chickering and Gamson, 1987), and ratings of student progress on 12 different objectives. The fact that a unique pattern of "most effective techniques" was found for each of the 12 objectives, and that the pattern was often different for classes of different sizes, underscores the complexity of trying to determine why instruction was effective or ineffective.

But identifying "strengths" and "weaknesses" by examining student ratings is not likely, by itself, to result in improved effectiveness. Centra (1993) observed that improvement occurs only when new knowledge valued by the teacher is acquired and then only when the teacher is motivated to change.[7] Cohen (1980) and Brinko (1991), in their reviews of the use of student ratings in improving teaching effectiveness, concluded that, without consultation, student feedback has no effect on improving teaching quality.

Although research identifying other factors related to student learning is less substantial, there is theoretical and experiential reason to consider a number of other characteristics (see, for example, Perry & Smart, 1997; Braskamp & Ory, 1994; Cohen & McKeachie, 1980). This includes adequacy/appropriateness of teaching objectives, quality of instructional materials, coherence of teaching strategy/plan, and comprehensiveness/timeliness of methods for appraising and reporting student progress.

While this paper is concerned with conducting formative evaluations, it does not address the question of how they should be used to stimulate improvement. Different plans have been employed by various institutions [e. g, Weimer (1990); Wright (1995); Centra (1993)]. In some, the department head/chair is expected to serve as "mentor" to the faculty, using evaluative feedback to provide suggestions for improving and monitoring efforts to implement these suggestions[8]. At others, a faculty development office has been established with one or more professionals who devote their time and energy to consulting with faculty members about ways to improve their instructional performances. Recently, "Peer Consultation Programs" have been developed which appear to be not only professionally credible but also acceptable to faculty (e. g., Hutchings, 1995; Sheppard, Johnson, and Leifer, 1998). A system designed by Bernstein (1996) pairs faculty members off in a series of mutual assistance activities focused on the three major elements of the instructional process—the syllabus (plan), classroom practices (implementation), and assessment of student achievement (outcomes).

Each of these approaches may be particularly advantageous on a given campus. What is clear from experience is that lasting improvements are almost never made without some kind of active assistance from another person.

[7]Motivation reflects both the faculty member's values/dispositions and the institution's priorities.

[8]Although there have been instances when this arrangement has been very successful, in general those charged with helping others succeed best when they have no responsibility for administrative decisions or recommendations affecting those they are serving.

## Evaluation Schedules

It is obvious that, for the evaluation of instruction to be usefully employed for either summative (administrative) or formative (improvement) purposes, it needs to be done with care and thoroughness. Demands made by the institution's instructional, research, and service programs are so intense that finding the time to conduct valid, comprehensive evaluations of these activities poses an important challenge.

The need for evaluation varies with its purposes and with the status of the faculty member. With respect to faculty status, evaluation needs are different for first year (non-tenured) faculty, other non-tenured faculty, and tenured faculty. These needs also vary depending on the purposes of the evaluation—formative or one of four types of summative. A realistic schedule for conducting evaluations of the faculty's contribution to the instructional program is suggested below.

First-year faculty. For many, this will be their first teaching experience. Most will have a number of special challenges (becoming acquainted with the characteristics and culture of a new institution and its undergraduate students; creating several new courses; adapting to a new environment). High energy and motivation levels will compensate for some of these difficulties, but first year results are not likely to be highly favorable. It seems appropriate to recommend that student ratings be collected for every course during the first term. On the basis of those findings, it may be possible to choose a focus for improvement efforts during the second term. At that time, colleague judgment might be used to supplement student ratings of the targeted course, while student ratings would again be collected for other courses.

Administrative decisions have to do with (1) retention and (2) salary recommendations. It will probably be necessary to make the latter before second term results are available[9]. Although the department chair/head will usually be able to make some relevant judgments on the basis of routine observations (conduct in faculty meetings; conscientiousness in placing book orders; submission of the teaching portfolio), it may be necessary to base recommendations for the instructional portion of the evaluation largely on student ratings. More information will be available for the retention decision, including student ratings for the second term, extended observations by the chair/head, and, hopefully, colleague ratings of the relevant features for one course.

Other non-tenured faculty. For institutions who employ a merit salary policy, an annual summative evaluation is required before making salary recommendations. Although such recommendations are important, they are less vital than those associated with tenure and rank, because the opportunity to correct salary-related errors is presented annually. The same is not true of the most central concerns with respect to non-tenured faculty members—the tenure decision (usually made after five or six years) and promotion. These decisions will largely determine the department's and institution's future quality and vitality. Therefore, they require evaluative information which is as valid and comprehensive as possible.

[9]In some institutions, participation in the merit salary increase program is limited to those who have been employed for more than one year. First year faculty may receive increases which approximate the average for the department, but are not evaluated for merit purposes because of the handicaps to performance cited in the text.

- It is suggested that, by the time a tenure recommendation must be made, at least two sets of student ratings (one fairly early in the faculty member's tenure; one relatively recent) be available
- for every course taught by the faculty member. In addition, at least two colleague reviews of each individual course ("early" and "recent") should be completed. It would also be prudent to collect at least two faculty ratings of "indirect contributions to the instructional program". Evaluations based on these data can also be used for supporting annual merit salary and reten-tion recommendations, but these will usually be less comprehensive than those used in making tenure and promotion recommendations.

Improvement efforts should be guided by formative evaluations of courses chosen by the faculty member on a schedule which is responsive to his/her situation. While such efforts are not a part of summative evaluations, recommendations with respect to tenure may well include a consideration of the degree to which such efforts have been made (as well as evidence that these efforts are succeeding); concrete engagements in faculty development are generally indicative of an individual who wants to remain "alive" and "vital" professionally.

mation than is needed for either summative or formative evaluation. But if the informational needs of the evaluation systems are known, the portfolio approach can almost always accommodate them.

The chart below illustrates one way to satisfy the needs for evaluating contributions to the instructional program without requiring an excessive commitment of time and energy. Modifications to this illustrative schedule will be required to reflect the special circumstances of individual institutions.

| Source of Information | Faculty Status | | |
|---|---|---|---|
| | 1st Year, Non-Tenured | Other Non-Tenured | Tenured |
| Student Ratings | Formative ratings collected for every course each term. | Formative ratings collected for every course on two occasions over a 5-year period. | Summative ratings collected for every course once every three years. Formative ratings collected as desired by professor. |
| Colleague Ratings of Specific Courses | Formative ratings collected for one course during the second term. | Formative ratings collected for one or two courses every year. | Summative ratings collected for every course once every three years. Formative ratings collected as requested by professor. |
| Colleague Ratings of Indirect Contributions | — | Ratings collected annually | Ratings collected every three years. |
| Department Chair Ratings | Observation, feedback after each term. | Annual observation and feedback. | Observation and feedback as needed. |

Tenured faculty. Except in rare cases, no further decisions about tenure and retention will be made. And for many, no further decisions about promotion will be made. In those instances where promotion is at issue, the evaluative evidence required includes that required for the tenure decision; in addition, evidence with respect to instructional versatility and the individual's qualifications for handling complex and difficult teaching assignments is relevant.

In institutions where merit salary policies are in place, all faculty members (including those with tenure) must be evaluated annually. With this exception, it is probably not necessary to conduct annual evaluations of such faculty members. On the other hand, a process for regularly monitoring their effectiveness in carrying out their assignments will help to identify "burnout" problems or those who, for a variety of reasons, have suffered a loss of enthusiasm for and commitment to their work. Therefore, it seems reasonable to collect student ratings of instructional effectiveness for every course the faculty member teaches at least once every three years. In addition, colleague reviews of each course should be made on an occasional basis, and ratings of indirect contributions should be collected at least every three years.

The process for collecting information relevant to instructional evaluation differs from campus to campus and from individual to individual. Perhaps the most dependable way to assure that all relevant information is available is the "teaching portfolio" (Seldin, 1997). Portfolios can be designed to support a number of purposes and may, therefore, contain more infor-

The question frequently arises as to how much weight should be placed on each source of evaluative information. Each institution should develop a policy which reflects the local relevance (importance, reliability, validity) of each. In general, if there is evidence supporting the value of the student rating process employed on the campus, results from it might account for 30–50% of the overall evaluation of instructional excellence. If the process for obtaining colleague ratings of specific courses is a thorough and conscientious one, results from it would normally merit about the same consideration as those from student ratings. However, since it is unrealistic to expect faculty to devote the time and energy required by annual assessments of every course, it seems reasonable to base about 25–35% of the overall evaluation on such ratings. If colleague ratings of indirect contributions are employed, they will probably be collected regularly for all faculty (except first-year faculty); it is suggested that these account for 10–15% of the overall evaluation. The department head/chair will also make annual judgments of contributions to the instructional program; these might also make up about 10–15% of the overall evaluation.

At individual institutions, the quality of evaluative information from various sources is likely to be uneven. The student rating program may, for example, be considered to be a sound one, but there may be no process for assessing the faculty member's *indirect* contributions to the instructional program; the chair's ratings may be based on documented evidence, but the system for gathering colleague ratings may lack credibility. To avoid giving undue weight to a given source of information, it is suggested that: (a) percentage figures be adopted which

represents the institution's judgment of the relative importance of each evaluation source *under ideal conditions*; (b) if information from a given source is unavailable, all faculty members be given a rating equal to the average of ratings from other sources; and (c) if information from a given source is believed to be of marginal validity, create two ratings for each faculty member—one using the "marginal" process, and one equal to the average rating for all faculty members from other sources. These suggestions are intended to ensure that the final evaluation figure is not unduly affected by a given source, regardless of how sound that source may be. They will inevitably reduce the degree to which evaluations differentiate between the "best" and "worst" teachers; but this is believed to be preferable to an over-reliance on single sources (student ratings; colleague ratings; etc.)[10].

## Summary

Although evaluation is inherently threatening to most faculty members, the vast majority take their assignments seriously and have a sincere desire to conduct them as effectively as possible. When the departmental environment is characterized by a strong commitment to mission, mutual respect and trust, and administrative support for faculty, a sound evaluation program can play a vital role in promoting both individual and organizational excellence.

Assessing faculty performance is a complex and time-consuming process. If it is done poorly or insensitively, it can have an adverse effect upon institutional quality. Whether or not individual institutions elect to commit the resources which valid evaluations require depends upon the degree to which they agree with three propositions:

1. All members of the institution should be accountable for their activities and performance.

2. The conduct and utilization of credible evaluation programs have an important influence on the welfare and future excellence of the individual, the department, and the institution.

3. When improvement efforts are supported by institutional policy and guided by comprehensive and valid appraisals of current functioning, the well-being of the individual and of the institution are positively affected.

[10]By employing "leveling" to deal with incomplete or unreliable data, the institution risks an inadvertent alteration of the priorities it assigns to "instruction", "research", and "service". For example, if ratings of instructional effectiveness differentiate only slightly among faculty, while ratings of "research" or "service" effectiveness vary widely, the latter will automatically have an increased impact on overall evaluations. Special procedures to protect against such unintended effects are needed.

IDEA Center
1615 Anderson Avenue
Manhattan, KS 66502-4073
Tel: 800.255.2757 or 785.532.5970
Fax: 785.532.5637
E-mail: IDEA@ksu.edu

©1999, IDEA Center

## References

Aleamoni, L. M. (1981). "Student Ratings of Instruction". In J. Millman (Ed.), *Handbook of teacher evaluation*, 110-145. Beverly Hills, CA: Sage.

Bernstein, Daniel (1996). "A Department System for Balancing the Development and Evaluation of College Teaching: A Commentary on Cavanagh". *Innovative Higher Education, 20*, No. 4, 241-247.

Braskamp, L. A. & Ory, J. C. (1994). *Assessing faculty work: Enhancing individual and institutional performance.* San Francisco: Jossey-Bass.

Brinko, K. T. (1991). "The Interactions of Teaching Improvement." In M. Theall & J. Franklin, (eds.), *Effective practices for improving teaching.* New Directions for Teaching and Learning, No. 48, San Francisco: Jossey-Bass.

Cashin, W. E. (1989). "Defining and Evaluating College Teaching", *IDEA Paper No. 21*, IDEA Center, Kansas State University.

Cashin, W. E. (1995). "Student Ratings of Teaching: The Research Revisited,", *IDEA Paper No. 32*, IDEA Center, Kansas State University.

Chickering, A. W. and Gamson, Z. (1987). "Seven Principles of Good Practice in Undergraduate Education". *American Association for Higher Education Bulletin, 39*, 3-7.

Cohen, P. A. (1980). "Effectiveness of Student Rating Feedback for Improving College Instruction: A Meta-Analysis of Findings." *Research in Higher Education, 13*, 321-341.

Cohen, P. A. & McKeachie, W. J. (1980). "The Role of Colleagues in the Evaluation of Teaching". *Improving College and University Teaching, 28*, 147-154.

DeZure, D. (1999). "Evaluating Teaching Through Peer Classroom Observation". In Seldin, P. & Associates, *Changing practices in evaluation teaching*, pp. 70-96, Bolton, MA: Anker.

Hutchings, P. (Ed.) (1995). *From idea to prototype: The peer review of teaching: A project handbook.* Washington, D. C.: American Association for Higher Education.

Keiz, L. & Waggoner, M. D. (1994). *Collaborative peer review: The role of faculty in improving college teaching.* ASHE-ERIC Higher Education Report, No. 2. Washington, D. C.: The George Washington University, Graduate School of Education and Human Development.

Seldin, P. & Associates (1993). *Successful use of teaching portfolios.* Bolton, MA: Anker.

Seldin, P. (1997). *The teaching portfolio: A practical guide to improved performance and promotion/tenure decisions* (2nd ed.). Bolton, MA: Anker.

Seldin, P. (1999). "Self-evaluation: What Works? What Doesn't?" In Seldin, P. & Associates, *Changing practices in evaluating teaching*, 97-113. Bolton, MA: Anker.

Sheppard, S. D., Johnson, M., & Leifer, L. (1998). "A Model of Peer and Student Involvement in Course Assessment." *ASEE Journal of Engineering Education, 87* (4), 349-354.

Weimer, Maryellen (1990). *Improving College Teaching.* San Francisco: Jossey-Bass, 1990.

Wright, W. Alan and Associates (1995). *Teaching Improvement Practices.* Bolton, MA: Anker Publishing Co.



BEST COPY AVAILABLE

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

# NOTICE

## Reproduction Basis

☑ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").